

17/09/2001

DE000294

1999 P 2739 WO  
PCT/DE00/02940

- 17 -

Patent Claims

1. A method for conditioning a database for automatic speech processing, in which with the aid of the database containing words in the form of graphemes and phonemes an assignment of the graphemes to the phonemes is set up automatically, comprising the following steps:
- a) assigning the graphemes to the phonemes of all the words which have the same number of graphemes and phonemes, the graphemes and phonemes being assigned to one another in pairs,
  - b) assigning the graphemes to the phonemes of all the words which have more graphemes than phonemes, all the graphemes firstly being assigned to the phonemes in pairs until an assignment error arises on the basis of the assignments determined hitherto, or there are present only at the end of the word one or more graphemes which are not assigned to a phoneme, and combining a plurality of graphemes to form a grapheme unit and assigning the grapheme unit to a phoneme, and
  - c) assigning the graphemes to the phonemes of all the words which have fewer graphemes than phonemes, a plurality of phonemes being combined to form a phoneme unit, and a single grapheme being assigned to them in such a way that the remaining grapheme/phoneme assignments of the word to be analyzed correspond to the assignments found under assignments),
  - d) assigning the words hitherto not assignable, the words being reexamined in terms of the phoneme units determined under c) and/or the grapheme units determined under b), and the graphemes being assigned to the phonemes while taking account of the phoneme units in case c) and the grapheme units in case b), and there being executed at least after step a) a correction step with the aid of which assignments of

AMENDED SHEET

17/09/2001

DE000294

1999 P 2739 WO  
PCT/DE00/02940

- 17a -

words which contradict the further assignments  
determined in step a) are erased.

2. The method as claimed in claim 1, characterized in

that after each of the steps a) to d) a correction step is executed with the aid of which assignments of words which contradict the further assignments determined in the respective steps or partial steps are erased.

5

3. The method as claimed in claim 1 or 2, characterized in that the correction step is executed on the basis of a static with which the frequency of the individual grapheme/phoneme assignments is acquired, the correction step being used to determine which assignments have a frequency which is below a predetermined threshold value and erases these assignments.

15

4. The method as claimed in claim 3, characterized in that the words which include an erased assignment are marked as not assigned so that they can be taken into account again in one of the subsequent steps.

20

5. The method as claimed in one of claims 1 to 4, characterized in that only vowels or consonants are combined to form a phoneme unit when assigning the graphemes to the phonemes of all the words which have fewer graphemes than phonemes.

25

6. The method as claimed in one of claims 1 to 4, characterized in that, when assigning the graphemes to the phonemes of all the words which have fewer graphemes than phonemes, vowels or consonants are firstly combined in a phoneme unit and assigned in accordance with step c), and if words which have fewer graphemes than phonemes remain impossible to assign, vowels are also combined with consonants

30

to form a phoneme unit and assigned in accordance with step c).

7. The method as claimed in one of claims 1 to 6,  
5 characterized in that when assigning the graphemes to the phonemes of all the words which have more graphemes than phonemes, the remaining grapheme or graphemes provided at the end of the word is combined to form a grapheme unit together with the last grapheme which has  
10 been assigned to the last phoneme of the word, and which is assigned to the last phoneme of the word.

8. The method as claimed in claim 7, characterized in that if an assignment contradiction results on the  
15 basis of the assignments determined hitherto, on the longest chain of graphemes that is free from assignment contradictions the next grapheme is combined with the last grapheme of the chain to form a grapheme unit, and a new attempt is made at the assignment, in which case,  
20 if it is impossible, again, to set up a grapheme/phoneme assignment, the next grapheme is combined with the grapheme unit last formed and a new attempt is made at the assignment.

25 9. The method as claimed in claim 8, characterized in that, if it is impossible to achieve a grapheme/phoneme assignment of a word with more graphemes than phonemes, the assignment is firstly started with a pairwise assignment beginning at the start of the word until an  
30 assignment contradiction results on the basis of the assignments determined hitherto, whereupon a pairwise assignment beginning at the end of the word is executed and, when only a single non-assigned phoneme is left over, the remaining graphemes are combined to form a  
35 grapheme unit and assigned to the one non-assigned phoneme.

10. A method for training a neural network for assigning graphemes to phonemes for automatic speech processing, characterized in that the neural network is trained with the aid of a database conditioned using  
5 the method of claims 1 to 9, the graphemes being input at input nodes, and the associated phonemes being input at an output node of the neural network.
11. A method for assigning graphemes to phonemes in  
10 the synthesization of speech, characterized in that the grapheme/phoneme assignment is executed by assigning an output pattern to an input pattern of the neural network trained using the method of claim 10, the input pattern comprising at least the letter to be assigned  
15 and, if present, at least one letter preceding in the word and one subsequent letter, and the output pattern having one phoneme.
12. The method as claimed in claim 11, characterized  
20 in that the input pattern comprises a plurality of letters preceding and subsequent to the letter to be assigned, it preferably comprising in each case three preceding and subsequent letters.
- 25 13. The method as claimed in claim 11 or 12, characterized in that the input pattern comprises the last output pattern.
14. The method as claimed in one of claims 11 to 13,  
30 characterized in

that the output pattern exhibits a grouping of the letters, that is to say the number of the letters combined to form a grapheme unit in the input pattern.

- 5 15. A method for assigning phonemes to graphemes in the recognition of speech, characterized in that the grapheme/phoneme assignment is executed by an assignment of an output pattern to an input pattern of the neural network trained using the method of claim
- 10 10, the input pattern having at least the phoneme to be assigned and, if present, a phoneme preceding in the word and a subsequent phoneme, and the output pattern having one letter.
- 15 16. A neural network for automatic speech processing, with the aid of which an assignment between graphemes and phonemes can be produced automatically, characterized in that the neural network has been trained in accordance with the method as claimed in
- 20 claim 10.
17. The neural network as claimed in claim 16, characterized in that it is stored on an electronically readable data medium.

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**